GPS – From Stones To Satellites Essay, Research Paper

The Stone AgeIn the early days of man, navigation was composed remembering objects as fixed points of reference. Leaving a trial of stones, marking trees and referencing mountains are examples of primitive navigational aids. The principles of this kind of navigation has evolved and is even present in today?s sophisticated navigational aids.

The Star Age (Trigonometry)Identifying points of reference on land was easy. However, man started to explore the oceans where the only points of reference were the sun, the moon and the stars. Naturally they became points of reference and celestial navigation began. The position of stars and their geometrical arrangement looks different at different positions on the earth. Therefore, by observing the configuration of stars, one could estimate his position. For better accuracy special optical instruments were invented to measure the angles between stars. The data was then transferred to special charts where a position could be calculated. However, measuring angles with stars was limited to clear nights and only accurate to several miles.

Radio Age (Distance = Velocity X Time)Around the middle of the century, scientists found a way to measure distances using radio signals. The basic concept is simple and works by using the relationship between distance, speed (around the speed of light for radio signals), and time. Accurate measurement of the signals time is important since on microsecond (one millionth of a second) equals 300m.

Below is how a radio receiver/transmitter (LORAN system) calculates ones position

The exact location of point A is programmed into a special radio receiver. The receiver when turned on measures the distance from the transmitter as 1000m. However, this does not tell us where we are but rather narrows our location to any point on a 1000m circle around the tower as shown in Figure 1.

Next assume that a second radio tower B is programmed into the receiver. The receiver when turned on measures the distance to tower B at 750m. Now we have two pieces of information: our distance to point A is1000m and our distance to point B is750m. So we are on circle A and circle B at the same time. Therefore we must be at the intersection of the two circles, one of the two points P or Q shown in Figure 2.

Measuring our distance to a third radio tower C would identify exactly where we are. Transmitters A, B, and C together are called a transmitter “chain”. A chain may have four or more transmitters in order to have better coverage. However, the range of a radio transmitter is about 500 kilometers and the system is only two dimensional, it can not calculate or attitude or elevation above an object.

GPS (Global Positioning System)Left the Stones and now into Satellites

GPS is similar to star based navigation. Except now the stars are man made and emit radio signals instead of light. GPS is composed of 24 satellites in 6 different planes located 11,000 miles above the earth that act as fixed reference points. By measuring the travel time of a signal transmitted from a satellites, a receiver can calculate its distance from that satellite. When receiving the signals from at least 4 satellites, a receiver can determine latitude, longitude, altitude, and time.

There are five master ground stations located at Hawaii, Ascension Island, Diego Garcia, Kwajalein, and Colorado Springs that continually track and correct the satellites for variations in position and time and errors.

Components of the SatelliteEach satellite is equipped with solar panels. These panels capture energy from the sun which provides power for the satellite throughout its life.

Internal components include such items as atomic clocks and radio transmitters. Each satellite contains four atomic clocks. These clocks are accurate to a billionth of a second or a nanosecond.

External components include a variety of antennas. The signals generated by the radio transmitter are sent to GPS receivers via the L-band antennas. Each of the 24 satellites transmits its own unique code in the signal. Since GPS was developed by the Department of National Defense, each satellite emits two separate signals, one for military purposes and one for civilian use.

How to determine the distance to a satellite How does the handheld receiver know when the signal left the satellite?

To better explain this lets use a simplistic analogy. Suppose there is a satellite orbiting 10 000km above you in space and you are on the ground with a hand held receiver. At the same time, the receiver and the satellite begin playing a song (The inky dinky spider). Than standing at the receiver, you would hear two versions out of sync, one from the receiver and one from the satellite (delayed). If you wanted to see just how delayed the satellite’s version was, you could start delaying the receiver’s version until they fell into perfect sync. That delay time is equal to the travel time of the satellites version. So we just multiply that time times the speed of light and you ve got the distance to the satellite. The song played is a complex digital binary (on/off) wave known as the Pseudo-Random Code.

Picture from http://www.casecorp.com/agricultural/afs/howgpsworks.htmlA graphical version of the Pseudo-Random Code (a series of on/off poises). Note the timing difference due to travel time of the satellite signal.

How to make the clock in the receiver an Atomic Clock (Clock Bias)Getting the time exactly right is very important in GPS. Errors in time are errors in distance. If the receiver’s clock is a thousandth of a second off, it would equal about 200 miles off course. To solve for this error a 4th satellite is used to coordinate time by comparing the travel time in 3 satellites to the travel time of the 4th satellite. If out of sync, the receiver will be detecting its position in two different places on the earth. By comparing these two positions, the receiver can adjust its internal clock to match the atomic clock in the satellite and thus positioning the receiver in one location.

Adjusting for clock bias is how GPS can not only tell you right where you are, but also exactly what time it is. It does this with the accuracy of an atomic clock but without the expense. This accuracy is useful in many areas, such as the timing of financial transactions. Wherever schedules need to be synchronized precisely, GPS can help.

Other ErrorsSA or Selective availability is the cause of the largest error for civilian users. It is created intentionally by the Department of National Defense so that our enemies cannot use GPS accurately against us. The Department of National Defense alters each satellite’s clock and orbital information by different amounts to degrade the accuracy. Military users have an encryption key so they can get to the accurate information.

Muitipath Errors GPS signals do not always take the most direct route to the receivers. They may reflect off of tall objects nearby, creating an echo. The signals are slightly delayed by the longer route they took, causing a multipath effect. Usually, receivers compensate for multipath errors by only accepting the first signal that reaches them, and by blocking out, signals that originate near the horizon, instead of overhead. A sign of mulitpathing would be when your position is moving/changing but you are stationary.

Atmospheric Refraction The greatest source of error, after Selective Availability, is the distortion of the satellites’ signals as they pass through the atmosphere. Out in space there is nothing to stow them down, but once they hit the earth’s atmosphere, the signals are bent (refracted), due to changing densities.

One would think that the earth’s troposphere, where all the weather is, would be the biggest problem, but in reality it accounts for only a small part of the error. Most of the trouble comes from higher up, in the earth’s ionosphere. The ionosphere is full of charged particles or ions that bend and slow down the signals. The density of these ions can be partially predicted because it fluctuates in daily and seasonal patterns. GPS uses a model to predict atmospheric conditions that can eliminate about half of the errors due to the ionosphere and troposphere. The satellites send information about atmospheric conditions in their navigation message for the receivers to use.

More on the IonosphereThe ionosphere, known as the electrically-charged region of the upper atmosphere, is significant in our technological lifestyle; it can interfere with or help radio broadcasts and GPS navigation. But much remains unknown about it. For researchers, the ionosphere is most interesting during sunspot-maxima, about every 11 years. The next maximum is due in 2000-2001. This is when the sun?s emission of X and UV rays are at its max. During this time a constellation of satellites known as FedSat will be in orbit and, measuring the changing ionospheres properties. Researchers at La Trobe university will use FedSat data, and computer imaging tricks borrowed from medicine, to build 3D pictures of the southern ionosphere. The gathered information will be used to improve satellite and radio emission broadcasting, including GPS.

Differential GPSDifferential GPS is a way to make GPS more accurate. It works by canceling out most of the natural and man made errors that come into GPS measurements. The process involves two receivers. The rover receiver (the one you have in your hand) and a base receiver (a fixed reference station). The base receiver through the use of precise radar to keeps track of the satellites exact position and altitude and from this computers develop a theoretical model (free of errors) for the signals travel time and location. The station than analyzes the incoming signal and compares it to the model. The difference between the two calculations is an “error correction factor?. Since the base receiver has no way of knowing which of the many available satellites a roving receiver might be using to calculate its position, the base receiver quickly runs through all the visible satellites and computes each of their errors. Then it encodes this information into a standard format and transmits it to the roving receivers via separate radio signals. So the roving receiver can use the data to correct its measurements.

Below is a diagram on how differential GPS works for ships. Courtesy of http://www.Colorado.EDU/geography/gcraft/notes/gps/gps.html